A UNIQUE MOMENT IN TIME:
CREATING THE NEXT GENERATION
COMPREHENSIVE ASSESSMENT SYSTEMS
TO SERVE ALL OUR CHILDREN WELL

By Pascal (Pat) D. Forgione, Jr., Ph.D.
Introduction

- Setting A Context: Connecticut’s Innovations in Assessment
- A Look Back: Standards-Based Reform Phase I
- The New Paradigm: Standards-Based Reform Phase II
- The Proposed Multi-State Assessment Systems for 2014-15
- The Imperative of Connecticut’s Continued Leadership
Connecticut’s Rich Tradition of Innovations in Assessment

EARLY BUILDING BLOCKS: Commissioner Shedd

♦ Connecticut Assessment of Educational Progress (CAEP) Program
  ▸ No Stakes Testing
  ▸ Broad Range of Subjects Assessed
  ▸ A Breath of Curricular Information Generated
  ▸ Irregular Cycle of Testing (Once Every Three to Five Years Per Subject)

♦ Educational Evaluation and Remedial Assistance (EERA) Program
  ▸ Low Stakes Testing
  ▸ Minimum Competency Testing Focus in Math, Reading and Writing
  ▸ Innovative Reading Measure (DRP) and Writing (Student Sample)
  ▸ Regular Cycle of Testing (Every Spring in Grades 3, 5, 7 and 9)
Connecticut’s Rich Tradition of Innovations in Assessment (Con’t)

ADVANCING THE FIELD OF ASSESSMENT: Commissioner Tirozzi

♦ Connecticut Mastery Test (CMT) Program

- Low Stakes Testing (Following the release of the *Nation at Risk* Report, 1983)
- Focused on Criterion-Reference Testing Design (Diagnostic Information in Math, Reading and Writing)
- Aligned to State Curricular Expectations, With Local Instruction and Professional Development
- Regular Cycle of Testing (Every Fall in Grades 4, 6 and 8 Originally; and Later Grade 10 and Science)

♦ Connecticut Teacher Assessment Program

- High Stakes Licensure Testing of Basic Skills and Content Knowledge
- Innovations Aligned to Teacher Education Program Reform (BEST)
- Established National Model for Measuring Pedagogical/Content Knowledge and Adopted by the National Board for Professional Teaching Standards
CT has been a national leader in standards-based education reform and pre-K – 12 assessment reform.

- 1997: report by the National Commission on Teaching and America’s Future highlighted CT’s investments in teacher quality, and subsequent gains in student achievement.
- 2008: State Board of Education’s Position Statement recognized “science as inquiry” as an organizing principle for science programs.
- 2009: Voluntary Connecticut Benchmark Assessment System (CBAS) to assess high-priority GLEs and provide immediate feedback.

A Look Back: Standards Based Reform

Phase I

♦ 1990’s

- Each state developed their own content standards, performance standards, and assessments
- State legislatures placed “high stakes” on assessments

♦ 2001: NCLB passed

- Required universal proficiency by 2013-14
- Significantly increased the amount of testing, and State costs
Standards Based Reform Phase I, continued

Phase I Problems:

- Standards vary by state, and often are “too many, too low”
- Proficiency-based system is blind to progress of students and schools at either end of spectrum
- Tests gained “center stage” rather than instruction
- Financial strain on states
- Improvements in achievement were inadequate, given the need

“The Proficiency Illusion,” The Fordham Institute, 2007

Figure 6 – Average ranking of states according to the difficulty of their mathematics proficiency cut scores across all grades (higher average ranks = more difficult standards)

Note: This figure shows the average rank in math across all grades measured within a state, where a high rank denoted a high proficiency cut score. Colorado’s math cut scores had the lowest average rank, while South Carolina’s cut scores had the highest average rank.
The Uniqueness of this Moment

Thomas Friedman in “The World is Flat” points out the importance of “inflection points” in history.

Concerning the launch of Netscape:

“We went from a world where value was created in vertical silos of command and control to one in which value is created horizontally on this platform by who you connect and collaborate with.

I would argue that shift from command-and-control to connect-and-collaborate is the mother of all inflection points. ... It is the biggest event, I would argue, to change human beings and how they interact, since Guttenberg invented the printing press.”

Thomas Friedman, 2010
The Common Core State Standards (CCSS) can become an “inflection point” for American public education - establishing a common foundation for building excellence and equity for all students.

The next critical steps:

- Developing common performance standards and assessment systems as part of the common foundation.
- Developing an interactive, online digital library that accelerates sharing of best resources, tools and practices.
RTTT Assessment Requirements for Comprehensive Systems

Requirements within the RTTT Assessment Program: $350 Million

♦ Building upon shared standards for college- and career-readiness

♦ Measure individual growth as well as proficiency

♦ Measure the extent to which each student is on track, at each grade level tested, toward college or career readiness by the time of high school completion and;

♦ Provide information that is useful in informing:
  - Teaching, learning, and program improvement;
  - Determinations of school effectiveness;
  - Determinations of principal and teacher effectiveness for use in evaluations and the provision of support to teachers and principals; and
  - Determinations of individual student college and career readiness, such as determinations made for high school exit decisions, college course placement to credit-bearing classes, or college entrance

The Assessment Proposals

The Process:

♦ Proposals were due from multi-state consortia on June 23, 2010
♦ Federal review and awards are anticipated by September, 2010
♦ New Consortium tests to replace current state NCLB tests by 2014-2015 school year

The Proposals:

♦ Two Comprehensive Assessment System Proposals:
  ▶ Partnership for Assessment of Readiness for College and Careers (PARCC, 31 states)
  ▶ SMARTER Balanced Assessment Consortium (SBAC, 26 states)

♦ One High School Assessment System Proposal:
  ▶ State Consortium on Board Examination Systems (SCOBES, 11 states)
The Comprehensive Consortia

PARCC State
SBAC State
Both consortia

June 23, 2010
Smarter Balanced (SBAC)

English Language Arts and Mathematics
Grades 3 – 8 and High School
English Language Arts and Mathematics
Grades 3 – 11
Noteworthy Features

♦ Advanced uses of technology

▶ In delivery:
  ▶ SBAC: computer adaptive delivery of EOY component
  ▶ PARCC: online, but fixed-form, at least initially

▶ Within test items to assess the CCSS college/career readiness skills

Mathematics Standard, High School:

Graph functions expressed symbolically … using technology for more complicated cases. … Graph exponential and logarithmic functions, showing intercepts and end behavior, and trigonometric functions, showing period, midline, and amplitude.

Writing Standard, Grades 6 - 12:

Gather relevant information from multiple print and digital sources, using search terms effectively; assess the credibility and accuracy of each source; and quote or paraphrase the data and conclusions of others while avoiding plagiarism and following a standard format for citation.
A successful examinee must enter the correct search terms and combine them using Boolean operators and syntax… and must also select other appropriate parameters for the search.

Examinees are scored on the characteristics of their searches as well as their ability to modify their search strategy in response to feedback.
In light of the initial feedback, the examinee performs a second search that
· includes the key content words, and
· properly limits and expands the search.

The feedback becomes increasingly specific with each unsuccessful search. After either a successful search or three unsuccessful searches, the task proceeds to the next step.
In scoring systems:
- Both SBAC and PARCC use electronic and artificial intelligence scoring.
- For items requiring human scoring, both will build electronic systems to provide training to scorers, distribute student response files to scorers remotely, and moderate the scoring process, for faster turnaround.

Digital Libraries: Both SBAC and PARCC include a digital library of released items, formative assessments, model curriculum frameworks and instructional units, professional development materials, student and educator tutorials and practice tests, etc.

Through-course summative assessments: PARCC proposes three through-course assessment windows per year for ELA and math. SBAC does not have through-course as a feature of the primary system, but plans to investigate a through-course alternative.
RTTT Assessment Requirements for High School Course Assessment Programs

Requirements within the RTTT Program:

♦ Develop new or adapted course assessments aligned to shared standards for college- and career-readiness

♦ Measure the knowledge and skills within a common set of college- and career-ready standards;

♦ Measure both achievement and growth

♦ Provide information that is useful in informing:
  ▸ Teaching, learning, and program improvement;
  ▸ Determinations of principal and teacher effectiveness and development and support needs; and

♦ Includes a process for certifying the rigor of each assessment and that courses covering similar content have common expectations for rigor.

♦ Two Competitive Priorities: a) STEM programs and b) technical certification programs.

The High School Course Assessment Proposal, SCOBES

English, Mathematics, Science, History and the Arts for High School Students
## Implementation Milestones

<table>
<thead>
<tr>
<th>SBAC</th>
<th>SCOBES</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2010 – 2012</strong></td>
<td>Item and task development and small-scale piloting</td>
</tr>
<tr>
<td><strong>2013</strong></td>
<td>Adopt common ELL and SWD accommodation policies</td>
</tr>
<tr>
<td><strong>Spring 2014</strong></td>
<td>Full field test of all items</td>
</tr>
<tr>
<td><strong>Summer 2014</strong></td>
<td>Adopt common Performance Level descriptors and develop achievement standards</td>
</tr>
<tr>
<td><strong>2014-15</strong></td>
<td>New Summative Assessments in Use</td>
</tr>
<tr>
<td><strong>Summer 2015</strong></td>
<td>Validation study of cut scores</td>
</tr>
<tr>
<td><strong>August 2015</strong></td>
<td>Adoption of common achievement standards</td>
</tr>
<tr>
<td><strong>January 2011</strong></td>
<td>Complete competitive selection of Board Exam Providers, aligned to Common Core</td>
</tr>
<tr>
<td><strong>May 2011</strong></td>
<td>Determine college-ready performance levels in English and Mathematics</td>
</tr>
<tr>
<td><strong>June 2011</strong></td>
<td>Establish “Move-On-When-Ready” cut score standard and college acceptance</td>
</tr>
<tr>
<td><strong>September 2011</strong></td>
<td>Begin Year 1 of implementation, grades 9 and 11</td>
</tr>
<tr>
<td><strong>August 2011</strong></td>
<td>Disseminate Year 1 results</td>
</tr>
<tr>
<td><strong>September 2012</strong></td>
<td>Begin Year 2 of implementation, grades 9 through 12, at expanded sites</td>
</tr>
<tr>
<td><strong>September 2013</strong></td>
<td>Begin Year 3 of implementation, grades 9 through 12, at expanded sites</td>
</tr>
</tbody>
</table>
Will This Become an Inflection Point?

Benefits:

✦ Fewer, clearer, next-generation standards

✦ Systems of high quality, aligned summative and formative assessments, with rapid results to teachers

✦ Online digital libraries for sharing of best resources, tools, etc.

✦ Leveraging of human and financial capital across states
Will This Become an Inflection Point?

Challenges

♦ Measurement challenges:
  - Measuring individual growth and “on track”
  - Use of individual student growth in determinations of teacher and principal effectiveness
  - Equating of through-course assessments and aggregating the results

♦ Curricular Flexibility at the Local level:
  - SCOBES offering choice of curricula
  - Controversial within SBAC and PARCC – what’s the right balance?

♦ ESEA Reauthorization: Will it align?

♦ Political Will: What about states that don’t win State RTTT grants? Changes in Governors?
The Imperative of Connecticut’s Continued Leadership

- Strong history of keeping the focus on quality teaching, and building the systems and supports needed.

- During SBR-II, Connecticut will be needed as a leader:
  - Demonstrating how to maintain that critical focus
  - Informing the development of high quality interim assessment systems
  - Leading in the sharing of best practices and quality resources
  - Sharing the excellent work in science inquiry and assessment
In Closing

The CCSS create an unprecedented opportunity to come together around common expectations, share our expertise with one another, and significantly improve teaching and learning for all students.
Contact:

Pascal (Pat) D. Forgione, Jr., Ph.D.
Distinguished Presidential Scholar and Executive Director
Center for K-12 Assessment and Performance Management
Educational Testing Service
823 Congress Avenue, Suite 816
Austin, TX 78701
E-Mail: pdforgione@k12center.org

www.k12center.org
Overview

The Center was established in 2009 to act as an independent catalyst and resource for the improvement of methodologies, technologies, policies and practices in K–12 assessment.

State initiatives to use common core standards and develop the next generation of common assessments, as part of the federally funded Race to the Top Assessment Program, open up an unprecedented opportunity for the development of integrated systems. This in turn would enable the development of assessments and reporting tools to support higher student achievement and equity.

This opportunity, however, also highlights the need for significant advances in several areas of measurement and assessment design, including:

- Use of through-course assessments within accountability systems
- Development of interactive computer-based items, including complex simulations and technology-based tools, to expand the range and complexity of skills that can be assessed and electronically scored
- Development of more valid and fair content assessments for English language learners
- Measurement of student growth and attribution of responsibility in teacher and principal evaluation systems

The mission of the new Center for K–12 Assessment & Performance Management (K–12 Center) is to create timely events where conversations regarding new assessment challenges and developments can take place, and to make available the best thinking and research on the range of issues facing national, state and local decision makers. Interested partners are encouraged to contact us.

To receive notices of the Center’s events and resources, please send an e-mail to Sharon Saldana at: ssaldana@k12center.org.
Resources Now Available on Our Website

**Next Generation Assessment Systems:**
- A 22-page report about the opportunities and challenges within the Race to the Top assessment agenda, based on the National Conference on Next Generation Assessment Systems held March 8 – 9, 2010 in Washington, D.C., with summaries and comparisons of four assessment system designs by prominent education and measurement experts:
  - Linda Darling-Hammond and Ray Pecheone, Stanford University
  - Steve Lazer, Educational Testing Service (ETS)
  - Lauren Resnick, University of Pittsburgh and Larry Berger, Wireless Generation
  - Marc Tucker, National Center on Education and the Economy
- Videos of the above presentations
- Brief summaries of the assessment system models as well as the full papers and PowerPoint presentations
- A video of the panel discussion involving Governor Roy Romer, former Governor of Colorado and Superintendent of Los Angeles Unified School District, Checker Finn, President of the Fordham Foundation and Jon Schnur, CEO of New Leaders for New Schools, about the uniqueness and urgency of this moment in American history
- Results of the pre-conference survey of the policy objectives that should drive the design of these future systems

**Measuring Individual Student Growth:**
- Two papers by leading measurement experts with alternative approaches to the measurement of individual growth:
  - Damian Betebenner, National Center for the Improvement of Educational Assessment and Robert Linn, University of Colorado
  - James Pellegrino, University of Illinois, Chicago
- Commentary: Wendy Yen, ETS and Michael Kane, ETS

**Using Student Growth Data in Determinations of Teacher Effectiveness:**
- Three papers on the challenges of using student growth data in determinations of effectiveness:
  - Edward Haertel, Stanford University
  - Henry Braun, Boston College
  - Robert Meyers, Value-Added Research Center, University of Wisconsin, Madison
- Commentary: Steve Lazer, ETS

**The Policy Issues Involved in the Measurement and Use of Student Gains:**
- One paper discussing some of the unintended consequences of high-stakes assessments and recommendations for future initiatives:
  - Dan Koretz, Harvard University
- Commentary: Robert Linn, University of Colorado and Drew Gitomer, ETS

**Designing Assessments to Inform Instruction:**
- Two papers offering new approaches to designing assessments for the purpose of informing instruction:
  - Margaret Heritage, University of California, Los Angeles
  - Mark Wilson, University of California, Berkeley
- Commentary: Randy E. Bennett, ETS and Cheryl Krehbiel, Washington, D.C., Public Schools

To learn more about the Center or download publications, visit us at [www.k12center.org](http://www.k12center.org)
Noteworthy New Features

Online Testing: Both of these Consortia are committed to moving quickly to the use of online testing. The advantages pointed out by the Consortia include:

• Reduced cost;
• Faster turnaround of scores to teachers, students, and parents;
• The option to utilize adaptive test delivery, which yields more precise scores for high- and low-achieving students;
• The ability to assess college- and career-readiness skills that involve the use of technologies, such as spreadsheets, and to assess skills and problem-solving processes that cannot be assessed on paper through items such as multi-step simulations; and
• A greater range of accommodations for students with disabilities.

Also, each proposal presses for rapid advances in the technologies for test items and electronic scoring of items. The Common Core State Standards contain some expectations that are difficult, if not impossible, to assess via traditional item types. The proposals call for aggressive research and development in the area of interactive assessment items that require students to move through multi-step simulations to solve complex problems. In addition, both proposals call for use of artificial intelligence in the scoring of additional item types in order to reduce cost and speed the return of results.

Informing Instruction and Program Improvement Throughout the School Year: Both of the Consortia plan to develop and provide to states and districts an online digital library of resources to support improved teaching and learning. Educators across the states could both access and contribute to this library of resources. Formative assessments, model instructional units, and professional development materials—all aligned to the shared standards—would be collected and shared.

Through-Course Summative Assessments: The RTTT Assessment Program application defines a through-course summative assessment as “an assessment system component or set of assessment system components that is administered periodically during the academic year.” The PARCC Consortium has included through-course assessments within its design. At the end of the first, second, and third quarter of the school year, students would complete short summative assessments that target a small number of essential skills or concepts. Research would be completed during the pilot phase to determine if and how the scores from these components could be aggregated with the end-of-year assessment for the final annual student score.

The SBAC proposal schedules all summative assessment to occur during the final 12 weeks of the school year. They also propose, however, to investigate the development of a through-course assessment system, which would replace the end-of-year test with several modules taken during the school year. Again, research would be conducted to verify the psychometric quality of this design before use.

Next Steps

The U.S. Department of Education will hand out awards by the end of September 2010. They may approve one or both of these proposals and may negotiate with Consortium leaders regarding the final designs. Decisions and final, approved designs will be available at: www2.ed.gov/programs/racetothetop-assessment/index.html

For more information about the Center for K – 12 Assessment & Performance Management, visit us at

www.k12center.org

Enhancing Assessments to Support Teaching and Learning

Next Generation Assessment Systems Proposed Under the Race to the Top Program

Prepared by the Center for K – 12 Assessment & Performance Management

2010
The Race to the Top (RTTT) Assessment Program will provide $350 million in competitive grants to support the development of a new generation of multi-state assessment systems. Unlike most existing state assessment systems, however, the vision expressed within the RTTT program features an integrated set of formative assessments for use by teachers within the flow of instruction, interim assessments to be given as progress checks throughout the year, and more focused summative accountability assessments (U.S. Department of Education, 2009).

The Assessment Program application calls for the development of new systems that will:

- Go beyond the current measurement of status/proficiency and measure individual student growth, for all students across the achievement spectrum;
- Measure the extent to which each student is on track, at each grade level tested, toward college- or career-readiness by the time of high school completion;
- Provide information that is useful in informing:
  - Teaching, learning, and program improvement;
  - Determinations of school effectiveness;
  - Determinations of principal and teacher effectiveness for use in evaluations and the provision of support to teachers and principals; and
  - Determinations of individual student college- and career-readiness, as such determinations made for high school exit decisions, college course placement in credit-bearing classes, or college entrance.

The Assessment Program invited applications by Consortia of states for a Comprehensive Assessment Systems for grades three – eight, and at least one grade in high school, and b) High School Course Assessment Programs. Applications were due by June 23, and two applications were submitted for the Comprehensive Assessment System grants:

- The Partnership for Readiness for College and Careers (PARCC), and
- The SMARTER Balanced Assessment Consortium (SBAC).

Awards will be handed out by the U.S. Department of Education by the end of September 2010. The new system(s) must be operational by the 2014 – 15 school year. The Center for Assessment, Development, and Accountability (CDA) is the lead agency to conduct an independent third-party review of the new multi-state assessment systems. The Partnership for the Assessment of Readiness for College and Careers (PARCC)

The purpose of this handout is to describe two of the proposals submitted to the U.S. Department of Education for the development of new multi-state assessment systems for grades three – eight and high school grades. The descriptions on pages 2 and 3 have been prepared by the Center for K – 12 Assessment & Performance Management and approved by Consortium representatives. For more detailed descriptions, we urge readers to refer to the full proposals at:

www.fldoe.org/parcc
www.k12.wa.us/SMARTER

The SMARTER Balanced Assessment Consortium (SBAC)

English Language Arts and Mathematics: Grades 3 – 11

The Partnership for the Assessment of Readiness for College and Careers (PARCC)

English Language Arts and Mathematics: Grades 3 – 11

Description of assessment system components:

- **Interim/benchmark assessments:** These optional computer adaptive assessments will provide near-immediate results on the same scale as the summative assessment. The item types will mirror the summative comprehensive assessment, but assess a smaller set of standards at a deeper level to provide more actionable diagnostic feedback. Reports will link teachers to appropriate formative strategies and professional development resources.
- **Performance tasks/events:** Students will complete 1 task in reading, 1 in writing, and 2 in mathematics annually, during a Consortium-defined testing window within the last 12 weeks of the school year.* Each task/event will require 1 to 2 class periods and will involve student-initiated planning, management of information and ideas, interaction with other materials and/or people, and production of an extended response such as an oral presentation, exhibit, product development, or an extended written piece. A combination of machine and teacher scoring will be used, with results returned within 2 weeks.*
- **EOY comprehensive assessment:** The EOY assessment will include approximately 40 to 65 questions per content area and will be presented to students using a computer adaptive assessment taken during the last 12 weeks* of the school year. It will include selected response, constructed response, and technology-enhanced items. A combination of immediate scoring by computer and teacher scoring using a distributed, moderated online scoring system will be used, and results will be returned within 2 weeks.* The system will support an additional opportunity for students, as locally determined.
- **Summative assessment:** The handout describes how both PARCC and SBAC will construct a summative assessment framework, the number of sessions/administered on any given day, and the timing of summative assessments.

*Time windows may be adjusted based on results from the research agenda and final implementation decisions.